

The Use of the Internet as an Engineering Resource

STEVEN A MCCLUNEY, MICHAEL W. HLAVINKA,
Bryan Research & Engineering, Inc.,
Bryan, Texas

ABSTRACT

A wealth of information covering almost every conceivable topic is available by way of the Internet. A significant amount of this information is useful to the engineer. Also, the Internet makes it possible to connect computers and communicate information in a variety of ways. This article discusses the basic aspects of the Internet as well as the ways in which an engineer might use of it.

Proceedings of the Seventy-Fifth GPA Annual Convention. Tulsa, OK: Gas Processors Association, 1996: 43-48.

[Bryan Research & Engineering, Inc.](#)

Visit our Engineering Resources page for more articles.

INTRODUCTION

Since its invention, the computer has been heralded as a way to make life easier, to process more data more accurately, and to make such data more readily available. Many of the claims have been fulfilled in amazing ways, whereas others are always just around the corner, waiting for the next level of hardware or software. In particular, we ask ourselves the following question: has the information age really made much more information available to us? Until fairly recently, this question would have to have been met with a resounding "No!", especially in regards to fields such as chemical or petroleum engineering. However, the advent of the Internet has been gradually changing this situation. As more data are placed at our electronic fingertips, it becomes increasingly important for people in technical professions to be familiar with the Internet, the vehicle by which we can communicate this information across the globe. Such knowledge can mean the difference between quickly ferreting out needed data or wasting valuable time wandering the seemingly endless maze of the net. Likewise, it is equally important to have a clear idea of what is and is not available on the network. This article explores the more popular tools of the Internet, and how they might be used by engineers as information resources.

What is the Internet?

Simply put, the Internet is a connection between your computer and all other computers that are connected to the Internet. This connection allows your computer to communicate, that is, to exchange data, with a remote computer. The form this data can take varies widely, and is closely related to the tool used to transfer it. It could be a short note to a coworker, or a data file of process parameters.

The Internet has its roots back in the late 1960's and '70s with the government funded creation of ARPANET. This network was conceived during the Cold War as a safeguard to communications in the event of a bomber attack. Due to the communication scheme envisioned and eventually implemented, the loss of one segment would not bring down the whole network. ARPANET rapidly became a means to quickly disseminate and exchange research data, although it was initially restricted to use by government and educational institutions. With time, more networks were linked and the restrictions on who could use the network were lifted. Although ARPANET no longer exists, its child, the Internet, has grown to link millions of computer systems around the world.

CONNECTING TO THE INTERNET

A number of options are available for connecting to the Internet, varying according to price and the amount of expected usage. First, one must decide whether one wants direct Internet connectivity versus connection to an information service such as CompuServe or Genie, which offers many additional features. Such services are usually reached over a standard telephone line by way of a modem, which somewhat limits the access speeds available. Also, these services generally do not give complete access to the Internet, although their additional features may make up for this. Most, however, do allow e-mail exchange with the Internet at the minimum, with some allowing file transfer protocol (FTP) and even web browsing.

The alternative to such an information service is a more direct connection via an Internet provider. This is a company that has a computer system connected directly to the Internet, and to which you are able to connect your computer in various ways. The least expensive method is to connect using a standard telephone line by way of a modem. In some cases, you are given an account which you log into with a terminal emulator (a program which simulates a terminal), at which point you have the same net access that you would have if you were actually sitting in front of a terminal connected directly to the provider's system. Such access is often limited by a text-only interface, which severely restricts the utility of the network. There is an alternate dial-in connection that makes your individual computer a node, or site, on the Internet. This scheme comes in two flavors: serial line Internet protocol (SLIP) and point-to-point protocol (PPP), with PPP being the more advanced (at least on a technical level), the user is unlikely to notice a difference between the two during operation. With this type of connection, you may use software written for your computer's graphical interface (such as Microsoft Windows). This allows Internet client software to have a much richer user interface.

If a high volume of network traffic is expected, it may not be cost effective to connect to the Internet provider via a dial-in scheme. Also, the dial-in connection is limited by the speed of the modem, which is ultimately limited by the transmission quality of a telephone line. With even moderate use, such speeds may not be enough, especially when transferring large files. Finally, there may be competition for your provider's phone lines, so that there could be times when you can't get access to the net. In this case, you should consider a dedicated line to your Internet provider. Such lines are often offered at a flat monthly rate, so that the cost is not dependent on usage. Also, you are generally able to get significantly better throughput with a dedicated connection than you would with a dialup connection (although, of course, the faster connections are priced accordingly). The choice of connection scheme is ultimately a factor of the size of your company and the amount of network usage you expect.

METHODS OF EXCHANGING INFORMATION ON THE INTERNET

There are a number of ways to send data over the Internet, many corresponding to particular tools used to create and send the data. Each of the most popular methods is discussed below:

FTP

The most basic method of exchanging data with a remote computer, file transfer protocol (FTP) allows you to transfer a file between your computer and another computer. This file could be a word processing document, a program, a data file, or any other file stored on the computer. The remote site must have an FTP server running (a program to handle incoming requests for file transfers), and you must log on to it (supply a recognized user name and a password) to be allowed access to the files on the remote system. Alternatively, your computer may act as the FTP server, allowing others to log on to it. This is particularly useful for sending a data file, for instance, from your computer to a coworker's computer, even if the coworker is halfway around the world. Many sites have elected to become clearinghouses for specific types of information, allowing anyone access. This is implemented by setting up an account with the user name anonymous (hence the term anonymous FTP), which anyone can log into and which only has access to specific files on the host computer. There are many anonymous FTP sites making available data on a multitude of topics. Although the large majority of available technical data is oriented toward computer applications, there are sites that provide information relevant to engineers. For example, the site <ftp://ftp.chem.utah.edu> (maintained by the chemistry department at the University of Utah) offers text file versions of material safety data sheets.

Telnet

Telnet allows you to log in and work on a remote computer system. It emulates a simple text terminal, usually one of the Digital VT series. You must generally have an account on the remote system, although there are sites which allow anonymous or guest telnet logins. Once connected, it is as though you were working on a terminal connected directly to the remote system. File modifications and program runs are performed on the remote computer, and a separate file transfer program such as ftp is required to retrieve any needed files.

At some sites, telnet may be used to access a service such as a database. For example, one may use telnet to access the Library of Congress database on books, patents, copyrights, and other registered information. Because telnet emulates a terminal, the database software running on the remote site is able to query you for required information, such as a search key.

Electronic Mail (E-mail)

E-mail is one of the most used methods of communicating via the network. At its best, E-mail offers a user-friendly way to compose and send text messages to other network users. Messages may also have files of arbitrary type attached, making E-mail a popular vehicle for transferring files. It is more limited in this respect than FTP, in that you may only send files (an inbound transfer must be initiated by the remote user). One should be aware that many mail servers restrict the length of E-mail messages, so that large files may not arrive intact. In these cases, FTP should be used to transfer the file. Also, in order to send a binary file, the file must be encoded (translated into an ascii, or text, file). Many E-mail clients are able to both encode files for attachment and to decode attached files. If your E-mail client is unable to do so, you would need a separate program for encoding/decoding.

E-mail requires both the sending and receiving computers to be running a mail server (or to be connected to a computer running a mail server). Messages are sent to a specific user, normally using a combination of their account name and the name of the computer where their account is located. For example, an E-mail address could look as follows: **Sheridan@candc.babylon5.gov** (refer to the discussion of Internet addressing below). E-mail capabilities and features vary widely, depending on the software used. Many packages offer features such as mailing lists, address books, message filing, spell and grammar checking.

News

Network news more closely resembles a physical bulletin board than any form of news from the media. The various parts that make up the network news system, collectively referred to as USENET, provide a public forum for exchanging messages. Using this system, you can carry on conversations, ask technical questions, and broadcast useful and relevant information to a group of people who have a common interest. More active newsgroups have hundreds of messages posted per day, with many more people who do not post reading the messages. Thus it is possible to interact with a wide variety of people.

Unlike E-mail, messages are posted to a newsgroup and may be read by anyone who subscribes to that group. Each newsgroup focuses on a specific topic. Newsreaders, the programs that allow you to access the posted messages, allow you to subscribe to the newsgroups of your choice (that is, to only view messages posted to these groups). Once subscribed, you can choose a group and browse through the messages that have been posted, and post your own messages/responses. Again, various news readers have widely ranging features and capabilities. Better ones allow threading (in essence, sorting the messages according to a subject line that was supplied by the poster, in order to group messages which are relevant to one another). Many newsreaders also incorporate a send-only E-mail client, so that you can respond to a post either in a follow-up post or directly by e-mail.

There are literally thousands of available newsgroups, covering nearly every conceivable topic. There are a number of chemistry and chemical engineering related groups, for example, such as sci.engr.chem. This newsgroup provides a forum for asking technical questions, for obtaining others' opinions (a current thread is asking "is it worth my while to take the PE exam?"), and for providing information which is known to be of interest to chemical engineers (for example, issuing a call for papers for an upcoming conference, or requesting sources of experimental data). USENET provides a very useful tool in that it can provide access to the experience and knowledge of a very large number of people, limited mainly by their willingness to help.

Archie/Gopher

Archie and Gopher provide mechanisms for searching for information available through the Internet. An Archie client (program) searches for specific filenames (or pieces of filenames) on anonymous FTP servers. For example, one could perform an Archie search for the substring "msds", and would find a number of anonymous FTP sites that provide material safety data sheets. Unfortunately, one would also find a host of other files whose names happen to contain the letters "msds". Archie is primarily useful if you happen to know the name of a file you are trying to find.

Gopher, on the other hand, is used to search for specific topics of information, and is not restricted to matching filenames. As such, a wide variety of information has been made available in a form that Gopher can browse.

World Wide Web

Perhaps the most significant computer development in recent years is the advent of the world wide web (or 'web' for short). The heart of the web is the web page, which is a file written in Hypertext Markup Language (HTML), a page formatting language. Web pages incorporate both text and graphics. One uses a web browser, such as Mosaic or Netscape, to view web pages over the net. Web pages also contain links, or hyperlinks, to other more or less related web pages, creating a vast, interconnected network. Web pages can be richly formatted with graphics and text, and HTML itself is constantly being extended by the producers of the major web browsers, with additions such as audio and movies. This creates an obvious vehicle for advertising, and many companies have put their product information on the web. The advantages to the company are low cost (if a company already has an Internet connection, publishing a web page may not add much to the cost) and wide viewership. Because of this incentive, the majority of technical information currently available through the web is in the form of product information and company contacts.

One cannot discuss the web without at least a short discussion of the web browser itself. In its most basic form, the web browser retrieves and displays web pages, and recognizes and follows links to other web pages. Today's more sophisticated browsers incorporate FTP (a link can point to a file, which will be downloaded to your computer when selected), E-mail (a link can point to an E-mail address, and selecting that link will open an E-mail client with that address configured as the recipient), and USENET news. As such, the web browser is arguably the single most important Internet client application. Also, as mentioned above, some of the companies that produce the web browsers have been modifying their browsers and HTML itself to greatly enhance the multimedia capabilities of web pages.

POPULAR NETWORK SOFTWARE

There are a multitude of quality software packages available for the various types of network access. In this discussion, we will confine ourselves to discussing packages which run under the Microsoft Windows environment, and are available via the Internet. Programs satisfying this last criterion are collectively known as shareware, although this term is actually specific to programs that can be freely downloaded and used for a trial period (usually 30 days) after which the software must be registered for a fee in order to continue use. Other distribution schemes include freeware for software that does not require a registration fee, and public domain for software that is not copyrighted and may be distributed, used, and modified at will.

There are far too many network programs to discuss in this forum, so only a few of the best (in the authors' opinions) will be presented. For a comprehensive listing of the available shareware network programs for MS Windows, and the locations for obtaining them, see **The Winsock Client Listing** web page.

E-mail

There are several excellent E-mail clients available. ConnectSoft's **E-mail Connection Internet Edition** is an only slightly watered-down version of their commercial product being offered at no charge, and has virtually all of the features wanted in a mail program with a polished user interface. **Pegasus Mail** is also a fairly complete freeware package and also has a reasonably good interface. Another popular client, **Eudora**, is somewhat more basic (in its freeware version) but is easy to configure and use.

FTP

Several FTP clients stand out from the pack. The first is John Junod's **WS FTP/WS FTP32**. This program is free for non-commercial use, shareware otherwise. Another innovative package is the shareware **CuteFTP**, which offers a different (and much improved, in this author's opinion) user interface over the traditional FTP client.

Telnet

There are a large number of Telnet applications available on the net. Most are adequate and have reasonable interfaces and prices. **Ewan** is a good freeware telnet client, for example. Other titles of note include **Trumpet Telnet** and **NetTerm**.

Web browsers

Until recently, there were relatively few choices in this area. Although a number of web browsers have become available, there are still only a few that dominate the arena. By far, the most popular web browser is **Netscape Navigator**, offering many cutting-edge features. An alternative is the freeware **NCSA Mosaic**, one of the oldest of the Windows web browsers, updated to be quite feature-rich though lagging slightly in the support of the newest extensions. Users of Windows 95 may freely download and use Microsoft's **Internet Explorer**, which supports many of the new HTML extensions (while adding a few of its own).

ADDRESSING ON THE INTERNET

It is appropriate at this point to discuss addressing on the net, i.e. how computers, users, web pages, and so on are uniquely identified to one another. Each computer connected to the Internet is assigned a unique Internet intercept protocol (IP) address, a set of four numbers ranging from 1 to 255 separated by periods. For example, the IP address for the anonymous FTP server at BR&E is 204.214.90.4, where 204.214.90 defines this address to apply to BR&E in general, and the trailing 4 identifies the specific computer. A site may be specified by the first number, the first two numbers, or the first three numbers (as above), depending on the site's network class. A small site has the first three numbers fixed, limiting it to less than 255 attached computers (the range of the fourth number). Although not the rule, it is possible for a single machine to have multiple IP addresses if it has multiple physical network connections.

The IP address is not particularly user-friendly, so a method was devised to assign more-or-less logical names to IP addresses: the *Domain Name System* (DNS). Each site is assigned a unique domain name which is based on the type of organization. For example, a company's domain name could be *example.com*, where the suffix *com* denotes a commercial organization. In the U.S. other suffixes include *edu* for educational institutions, *gov* for government agencies, *org* for non-commercial organizations, *mil* for military sites, and *net* for network providers or administrators. Sites outside of the U.S. do not conform to this standard; each country is assigned a unique suffix (for example, *ca* for Canada and *de* for Germany), but need not include the codes given above. The name preceding the period is usually (though not necessarily) taken from the company's name, if possible.

The domain name is appended to another name to specify an individual computer. For example, a computer named *filesaver* at the example site becomes *filesaver.example.com*. The name itself is actually arbitrary and more than one name may refer to a single computer (even for computers with a single network connection). The resolution of site names to IP addresses is performed on designated systems referred to as *domain name servers*.

Programs such as FTP and telnet require only an address or a computer name in order to connect initially (although both require a username and password afterward). E-mail, on the other hand, is targeted at a specific user on a given machine, so an E-mail address consists of two parts: the username and a site, separated by an @. On some systems (notably VAXs) additional information is required as part of the address, but this is the exception. For example, an E-mail address may look as follows: *webmaster@bre.com*. This routes the message to whoever has the account named *webmaster* at the site specified by the name *bre.com*. Note that the site (the part following the @) is not necessarily a machine or domain name – it may be a mail exchange, which serves to forward the message to a particular mail server. The mail exchange does not necessarily resolve to an IP address as it is merely a reference to a machine, rather than the name of the machine itself.

A web browser loads a specific file (i.e. the web page) at a site, and so this file name must be specified, although there are several default file names (such as welcome.html) that the host computer will check for if a name is not given. Because the browser is able to access files using various protocols, a prefix is added to the computer name/file name. For example, the web address for BR&E's home page is **http://www.bre.com/welcome.html**. The prefix "http://" tells the browser that the address is for a web page. This is followed by the name of the computer, *www.bre.com*. The last part is the file name of the web page to view. If this last part is omitted, *www.bre.com* will look for the default names (one of which is welcome.html) before reporting an error. The entire address, combining the prefix, site name, path and filename, is called a Uniform Resource Locator (URL). An example of an URL for a file other than a web page is **ftp://ftp.bre.com/pub/downloads/hardware.dat**. This tells the browser to attempt an anonymous FTP login to the computer ftp.bre.com, then to try to retrieve the file *hardware.dat* in the directory /pub/downloads. URLs may also specify E-mail addresses, Gopher searches, and other types of connection services on the world wide web.

INTERNET TECHNICAL DOCUMENTATION

The Internet is based on a variety of technical specifications. For example, there are specifications on how domain name servers that convert site names to IP addresses must function. All of these specifications are described in a group of Request for Comment (RFC) documents. Each RFC document is assigned a unique number. For the interested reader, these documents along with an index may be obtained by anonymous FTP or via the world wide web.

THE ENGINEER AND THE INTERNET

It is now time to focus on the question we asked at the outset: has the information age really made much more information available to us? The answer is 'yes', but it is still a qualified 'yes'. There are still several factors that limit the amount and types of available data that are relevant to the engineer. The first, and perhaps most important, are incentive and property rights. A company is not likely to spend time and money making technical information available without a clear return on the investment. Thus many companies have showy web pages with glowing product descriptions, but relatively few are offering hard numbers and technical specifications. Eventually this may change, as companies begin to see beyond the possibility of inexpensive advertising, and recognize the advantages of putting data into the hands of prospective customers.

Educational institutions, the other major source of technical information, are only slightly less concerned about advertising, although the distribution of nearly independent computer systems among various departments tends to mitigate this. Already, it is possible to find descriptions of ongoing research projects, and it seems likely to be only a matter of time before results are commonly presented on the net in much the same way as they are presented at technical conferences.

Newsgroups

As mentioned above, there are a number of newsgroups that are of interest to engineers. Some of these are briefly mentioned in Table I, with a focus on those relevant to chemical engineering. This group is far from complete, but should hit on the major sources of information.

Newsgroup	Topic Discussed
sci.engr.chem	Chemical engineering in general
sci.chem	Chemistry in general
sci.chem.electrochem	Electrochemistry/electrochemical engineering
sci.chem.electrochem.battery	Batteries and fuel cells
sci.chem.coatings	Coatings/platings
sci.engr	Engineering in general

sci.engr.control	Control theory (process and otherwise)
sci.engr.mech	Mechanical engineering in general
sci.engr.metallurgy	Metallurgy
sci.engr.safety	Safety in various engineering disciplines

Web Sites

Again, the lists of general web sites given in Table II and company web sites given in Table III, are not intended to be an exhaustive listing of web sites of interest to an engineer, but are rather intended to be samples of what is available, providing starting points through which you may begin to search the web. As such, several indices are listed in addition to the engineering- specific sites. These tables also focus primarily on web sites relevant to chemical engineering. Note that the sites listed are subject to change or withdrawal without notice.

Web address (URL)	Description
http://www.yahoo.com/Science/Engineering/Chemical_Engineering	Index of chemical engineering sites
http://www.che.ufl.edu/www-che/index.html	Index of chemical engineering sites
http://www.ccohs.ca/Resources/hshome.htm	Index of health/safety related sites
http://www.galstar.com/~gpa	Gas Processors Association
http://www.api.org	American Petroleum Institute
http://www.uoknor.edu/igupt	Institute for Gas Utilization and Processing Technologies
http://www.gri.org	Gas Research Institute
http://www.spe.org	Society of Petroleum Engineers
http://www.doe.gov	U.S. Department of Energy
http://www.epa.gov	U.S. Environmental Protection Agency
http://www.ansi.org	American National Standards Institute
http://www.astm.org	American Society for Testing and Materials
http://www.acs.org	American Chemical Society
http://www.che.ufl.edu/www-che/aiche	American Institute of Chemical Engineers
http://www.asme.org	American Society of Mechanical Engineers
gopher://gopher.sdsu.edu/11/SDSU%20Campus%20Topics/Departmental%20Information/Art%20Department/The%20Ceramics%20Gopher	Ceramics database
http://www.vuw.ac.nz/non-local/chemistry	A database of (non-network) chemical resources
http://www.st.rim.or.jp/~shw	Reagents chemical database
http://www.cchem.berkeley.edu/Table/index.html	Periodic table of the elements

Web address (URL)	Description
http://www.3m.com	3M
http://www.alliedsignal.com	Allied Signal, Inc.
http://www.amoco.com	Amoco Corporation
http://www.basf.com	BASF Corporation
http://www.bre.com	Bryan Research & Engineering
http://www.chevron.com	Chevron
http://www.dow.com	Dow
http://www.dupont.com	DuPont
http://www.eastman.com	Eastman Chemical Company
http://www.gcchem.com	Gulf Coast Chemical
http://www.hp.com	Hewlett-Packard
http://www.honeywell.com	Honeywell
http://www.mobil.com	Mobil
http://www.slb.com	Schlumberger

http://www.shellus.com	Shell
http://www.texaco.com	Texaco
http://www.york.com	York International

CONCLUSION

The Internet has made available a wealth of information and has opened many new avenues of communication. In the latter case, with electronic mail and the ability to send documents and data to coworkers around the globe, the benefits of an Internet connection are immediate and substantial. In the former case the net is still in its infancy, although it is growing daily. There is useful information to be found (among a myriad of less useful information), although most is oriented towards marketing. Technical product information is still a rarity, and as such should not be the basis for deciding to connect to the Internet.

REFERENCES

1. P.H. Salus, "Casting the Net", Addison-Wesley Publishing Company, New York, 1995.
2. S. Estrada, "Connecting to the Internet", O'Reilly & Associates, Sebastopol, California, 1993.
3. See, for example, world wide web site <http://www.ncsa.uiuc.edu/demoweb/html-primer.html>.
4. <http://www.mbnet.mb.ca/winter/index.html>. E. Sinkovits, 1996.
5. A. Marine, S. Kirkpatrick, V. Neou, and C. Ward, "Internet: Getting Started", Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1994.
6. Anonymous ftp: <ftp://ftp.internic.net>. World wide web: <http://www.internic.net>.

copyright 2001 Bryan Research & Engineering, Inc.